

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 02-118842

(43)Date of publication of application : 07.05.1990

(51)Int.Cl.

G06F 9/46

(21)Application number : 01-143937

(71)Applicant : DEMAX SOFTWARE INC

(22)Date of filing : 06.06.1989

(72)Inventor : ESBENSEN DANIEL MARK

(30)Priority

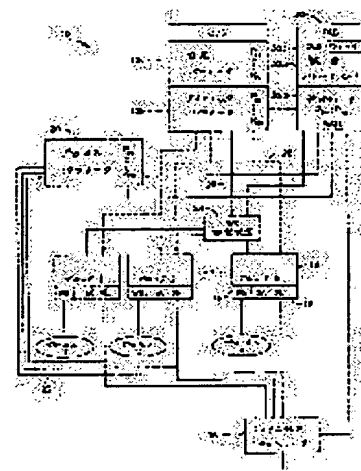
Priority number : 88 203227 Priority date : 06.06.1988 Priority country : US

(54) DYNAMIC LOAD BALANCE FOR COMPUTER TO BE USED BY MANY USERS

(57)Abstract:

PURPOSE: To improve the responsiveness of a system and to increase memories to be used by dynamically adjusting a system operation parameter in a multiple users computer system in accordance with a load or the system.

CONSTITUTION: The operation system parameter is used in respective process control blocks 14. An operation limit or a characteristic for respective processes active in the computer system are set. It is set by setting memory positions allocated to the respective processes at the time of initialization or an operation setting control value in the setting of a register 16. A WS control value is changed in accordance with a historical process characteristic for increasing an operation set page list for reducing a page accident or for reducing an operation set page list for memory re-allocation which is possible. Thus, the block 14 is connected to an operation control system 12 through a process adjusting device 24.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision]

of rejection]

[Date of requesting appeal against examiner's
decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2000 Japan Patent Office

⑫ 公開特許公報(A)

平2-118842✓

⑤Int.Cl.⁸

G 06 F 9/46

識別記号

3 4 0 D
F

庁内整理番号

7056-5B
7056-5B

⑬公開 平成2年(1990)5月7日

審査請求 未請求 請求項の数 15 (全22頁)

⑭発明の名称 多使用者コンピュータ用ダイナミック負荷平衡

⑯特 願 平1-143937

⑰出 願 平1(1989)6月6日

優先権主張 ⑱1988年6月6日⑲米国(US)⑳203,227

⑳発 明 者 ダニエル・マーク・エ アメリカ合衆国、カリフォルニア州 92130, サン・デイ
スベンセン エゴ、カーメル・ビュ 4075-10㉑出 願 人 デマックス・ソフトウ アメリカ合衆国、カリフォルニア州 94404, サン・マテ
エア・インコーポレー オ、シュート 500, ベイカー・ウエイ 999
テッド

㉒代 理 人 弁理士 鈴江 武彦 外3名

明 細 書

1. 発明の名称

多使用者コンピュータ用ダイナミック負
荷平衡

2. 特許請求の範囲

(1) 中央処理装置(CPU)と、其アドレスされる実上のメモリ動作システムと、システムリソースへのアクセスを割当てるための予め割当てられた制御値を有するシステム動作パラメータの固定されたセットとを使用し、リソースはCPU処理サイクルおよび内部メモリを備え、各プロセスに対する1組の測定された特性パラメータに対する値を発生するために1以上のモニタルーチンを使用する多使用者コンピュータシステムにおける、システム動作パラメータ値の調節によつてシステムリソースをダイナミックに再割当てする方法において、

メモリ番組位置に予め定められたセットの所望の特性値を与え、

第2のメモリ位置に各プロセスに対する前記測

定された特性値を記録することによつて各プロセスに対する現在のシステム応答特性を算出し、

各対応する測定された所望の特性値に対する予め定められた関係にしたがつて調整係数を発生し、

この調整係数に回答して前記システム動作パラメータを調整するステップを含むことを特徴とする方法。

(2) 予め定められた時間待機して前記累積ステップに戻り、前記待機の期間は前記調整係数に回答して複数の予め設定された長さの待機時間から選択される特許請求の範囲第1項記載の方法。

(3) それぞれ約6.30および90秒の長さの少なくとも3個の第1と第2と第3の予め定められた待機時間を有している特許請求の範囲第2項記載の方法。

(4) 前記コンピュータシステムは前記方法を終了したいときに終了命令を出力し、さらにこの終了命令に対するチェックと、それに対応する再割当て方法の終了と、そうでない場合の前記記録ステップへの復帰のステップを有している特許請求

の範囲第2項記載の方法。

(5) 前記予め割当てられたシステム制御値を検出し、この検出された値を第3のメモリ位置中に蓄積するステップを有している特許請求の範囲第1項記載の方法。

(6) 前記システム動作パラメータはシステム初期化において前記コンピュータシステムによつて自動的に負荷される特許請求の範囲第1項記載の方法。

(7) 前記動作パラメータは、中央プロセッサアクセス時間に対する予め選択された最小値および最大値、プロセッサ内部セット中のメモリの頁の最小数、内部セットインクレメントの大きさ、内部セットデクレメントの大きさ、最大および最小頁事故割合、およびDMA伝送のためのデータ伝送グループの大きさを制御するパラメータを含んでいる特許請求の範囲第1項記載の方法。

(8) 前記特性パラメータは、頁事故割合、頁事故、内部セットの大きさ、休止待機、頁衝突待機、および各プロセスに対する自由頁に対するバ

ラメータを含んでいる特許請求の範囲第1項記載の方法。

(9) 前記調整ステップは、

前記頁事故割合が高低限界を越えるプロセスに対して動作セットの大きさおよびアクセス時間を増加し、

前記頁事故割合が高低限界の間の範囲のプロセスに対して動作セットの大きさおよびアクセス時間を減少させるステップを含んでいる特許請求の範囲第8項記載の方法。

(10) 前記特性パラメータおよび関連する値を第3のメモリ蓄積位置に蓄積するステップを含んでいる特許請求の範囲第1項記載の方法。

(11) 前記第1と第2のセットの制御パラメータを蓄積した後にデータを蓄積するためにシステムバッファを初期化するステップを含んでいる特許請求の範囲第1項記載の方法。

(12) プロセスおよびシステム応答特性における統計的指標を編集するステップを含んでいる特許請求の範囲第1項記載の方法。

(13) 中央処理装置(CPU)と、このCPUに接続された頁アドレスされる実際上のメモリと、システムリソースへのアクセスを割当てるための予め割当てられた制御値を有するシステム動作パラメータの固定されたセットとを有し、リソースはCPU処理サイクルおよび内部メモリを備え、システムリソースの割当に関する各プロセスに対する1組の測定された特性パラメータに対する値を発生するために1以上の特性モニタを使用する多使用者コンピュータシステムにおけるリソースの割当てを制御する装置において、

前記CPUに接続されてシステム動作パラメータおよび対応するシステム制御値を蓄積するシステム蓄積手段と、

各プロセスに対する前記特性パラメータと関連する値を蓄積するために前記CPUに接続されている状態蓄積手段と、

前記CPUに接続されて動作し、各アクティブプロセスに対する前記特性値を測定しそれらを前記状態蓄積手段に蓄積するモニタ手段と、

前記モニタ手段に接続されて測定された特性値を受信し、1組のシステム状態値を形成するためにそれらを累積する状態累積手段と、

前記特性パラメータに対する1組の調整係数を提供するために前記CPUに接続されているリソース制御手段と、

この調整係数に反応して前記制御値を調整するために前記リソース手段および前記CPUに接続されている調整手段とを具備していることを特徴とする多使用者コンピュータシステムにおけるリソースの割当てを制御する装置。

(14) 前記調整手段および前記モニタ手段に接続されて前の調整値に応じてモニタおよび調整の周波数を決定するモード選択手段を備えている特許請求の範囲第13項記載の装置。

(15) 情報を蓄積するために少なくとも1組のバッファ手段を備えている特許請求の範囲第13項記載の装置。

3. 発明の詳細な説明

[産業上の利用分野]

この発明は、多使用者コンピュータシステムに関し、特にシステム負荷に応答してリソース割当てをダイナミックに調整することによってシステム応答性を改善し、利用できるメモリを増加させる方法に関する。この発明はさらに、特性フィードバックループおよびプロセス特定動作パラメータの調整によりコンピュータシステムリソースを管理し、動作効率を改善するための方法および装置に関する。

〔従来の技術〕

多使用者コンピュータシステムはたとえどんなに大きくても内部および外部メモリは限定された有限のメモリ量を有し、また個々の処理の実行或いは使用者の相互作用の処理のために利用できる中央処理装置(CPU)処理時間または1秒当りのサイクルの最大量を有している。典型的な多使用者状態では広い範囲に散らばった処理要求を有する多数の使用者が負荷状態の変更下のそれらの限定されたリソースに対して同時に結合することがしばしば生じる。それ故リソース管理システム

が、自然に発生する計画の衝突を最小にし、リソースにおける要求の衝突によるシステムの停頓または厳しい特性の劣化を阻止するためにそのようなシステムに設けられなければならない。

典型的な管理技術はシステムリソースに対する指定された処理受信プロンプトのアクセスを確保するための使用者またはプロセス優先度の設定である。この方法では、プロセスはそれらの出力または関連する使用者の相対的重要度に基づいて優先度を割当て、一般的に連続してより高い優先度が割当てられ、一方それらがリソースにアクセスするまで実行は待機される。しかしながらこの優先度による方法は遅延計画法に過ぎず、それらはシステム負荷を単に時間的にシフトまたは分配するだけで、そのリソースの効率または使用を改善するものではない。この方法はコンピュータシステムを遅延されたプロセスに応答するのを少なくし、同じ優先度レベルを有する多数の使用者に対するシステム応答特性を改善することは少ない。遅延計画は少数の高い優先度動作または相互作用

プログラムによって実行される多数のCPUおよび/またはメモリ集中動作の存在において限定された形態を有する。

システムリソースの管理の別の基本的技術は各種の動作タスクまたは機能のための動作システム制御パラメータの選定である。各制御パラメータはコンピュータシステムの特定の機能または動作、例えば使用者当りの最小および最大メモリ割当て、最小または最大CPUアクセス時間、またはI/Oアクセスと関連するフラグまたはラベルとして作用し、その機能または動作のための動作限界を設定する制御値を割当てられる。このようなパラメータの例は量子というラベルのパラメータであり、それは典型的にはコンピュータシステムに応じてCPUと内部メモリのいずれかに対するプロセスに対してに割当てられた最小アクセス時間を決定する。これは同じプロセッサに対する多重量子アクセス期間を与えるために優先度と組合わせて実行される。同時にCPUアクセス時間の最大限度はプロセスを周期的に待機状態に入らせ、他

のプロセスがアクセスすることを許すようにされる。付加的なパラメータのセットは各プロセスにより使用されるメモリの量について制限し、その量は増加または減少させることができる。

動作システムパラメータに対する値は一般的全体システムの要求、利用されるリソース、計画された作業負荷の考察後選定される。パラメータに割当てられた値はある方法で付勢され、或いはリセットされるときコンピュータの固定されたシステム初期化の一部としてシステムに自動的に組込まれる。パラメータ値は一定のままであり、システムの周期的同調のためにシステム管理者によって変更されるためにのみアクセス可能である。このような変更はシステム経験に基づいて、または使用者およびリソースの拡張に応じて負荷および使用者要求の変更を補償するために行われる。このため、1組の特性パラメータが典型的に設定され、内部モニタルーチンによるこれらのパラメータに対する累積された値はある変数を周期的に測定してボトルネックその他の問題を確実処理す

るのを助ける。

しかしながら、パラメータまたはそれらの関係する値は処理の実行を管理するための完全なテンプレートを与えるけれども、それらのパラメータは個々の、または特定のプロセス負荷および応答問題に適用し、或いはそれに焦点を絞る能力に欠けている。パラメータはまた迅速な、または短いダイナミック負荷変化には応答しない。

多使用者コンピュータシステム、特に事実上のメモリアドレスシステムにおけるメモリ割当ては一般に頁と呼ばれる予め定められた単位のメモリを使用する。各頁は転送され、それでなければ単位として動くメモリ位置のブロックまたはグループを表わす。コンピュータシステム動作パラメータ値は頁の全体数を設定し、各プロセスは所定の時間にプロセスに割当てられた内部頁の数と同様に使用される。プロセスが1以上の内部にない頁からのデータを必要とするとき、“頁事故”が生じ、そのデータは内部セット中の1以上の頁中のデータに対して交換される。

まず、作業または内部セットの大きさに対する調整は前述のように静的パラメータによつて制御される。内部セットの大きさにおける変化中に許容される頁またはインクレメントまたはデクレメントの最小および最大数は全て固定した数である。それ故そのような調整は激しく変化するダイナミックな問題に対する平均的解決を与え、個々の処理負荷条件における変化に回答することは不可能である。第2に、デクレメントステップよりも大きいインクレメントステップの使用によつて、調整処理はメモリの多過ぎおよび少過ぎるメモリ間の振動モードに入る傾向があり、過度のCPU時間が作業セット調整に消費される。コンピュータ製造業者は一般にそのような調整方法はそれらがリソース割当てに対してさらに問題を生成するために行われてはならず、使用されないものと教えている。

〔発明の解決すべき課題〕

必要なものは各種の負荷および処理状態下の多使用者コンピュータシステムの特性をモニタし、

もしもコンピュータシステムの能動処理の全てにわたるあまりにも多くの頁事故が生じるならば、CPUは入力/出力(I/O)管理の付加的時間を費やして内部メモリ位置と非内部メモリ位置との間で交換する。また内部メモリ中へおよび中からの頁取替えの作用はシステムの応答を遅くする磁気媒体に対する読取り書込みが必要である。この形式の通信または伝送管理のためのCPUサイクルの損失は他の計算および相互作用プロセスのために失われたCPUリソースの量に影響する。

ある別のコンピュータシステムでは、内部セットの外側の頁にアクセスする処理により経験された頁事故の数をもつと内部頁を保持するためにメモリを集中的に動作させるように内部セット限定に影響を与えるために使用できる。メモリ割当てを調整する動作システムの能力は長期間の周期的保守割合に対向するように処理中このリソースのさらに効率的な使用を可能にする。しかしながら、現在の内部セット調整方法はいくつかの問題を有する。

各プロセスに対するメモリおよびCPU時間割当てに対するダイナミック調整を行うための方法および装置である。使用される技術は振動を実質上生じないことが必要であり、非常にメモリ効率がよく、簡単に現在の動作システムと関連して構成することができる必要がある。その方法はまた所望のときにもとの動作特性に順序正しく戻らなければならない。その方法は大きな多使用者システムにおける補償またはサービス契約同意を危険にさらすような基本動作システムのどの部分における重ね書きもしてはならない。

技術における上記問題の観点から、この発明の目的は、各使用者またはプロセスに対するコンピュータリソースの利用を最良にするために多使用者または多プロセスコンピュータ中のシステム制御パラメータをダイナミックに調整する方法を提供することである。

この発明の利点はコンピュータシステムのリソースが高いダイナミック負荷でも適合するように周期的なインターバルで自動的に調整されること

である。

この発明の別の目的は、パラメータが個々のプロセスに対するリソースの割当てを最良にするためにプロセス制御ソフトウェアの制御パラメータを設定するようにダイナミックな調整動作を行う方法および装置を提供することである。

この発明の付加的な利点は、リソース割当てが振動特性を避けるように各プロセスに対して別々に調整されることである。

【課題解決のための手段】

これら、およびその他の目的および利点はシステムで行われる別々の処理によりシステム上の負荷の変化に応じて多使用者コンピュータシステム中のシステム動作パラメータをダイナミックに調整する方法によつて達成される。これは中央処理装置(CPU)と、アドレスされる実上のメモリと、各処理によるシステムリソースへのアクセスを割当てるための対応する制御値のセットと組合わされる予め定められたシステム制御パラメータとを使用する多使用者コンピュータシステム

数の予め設定された長さの待機時間から選択される。それぞれ約6、30および90秒程度の長さの3個の待機時間を有していることが好ましい。この方法のステップは終了命令が受信機されるまでは周期的に繰返される。コンピュータシステムが方法のステップの終了を要求されたかどうかを決定するチェックが周期的に行われる。

予め設定された動作システム制御パラメータ値は電力設定またはリセットにおいて生じる初期化においてコンピュータシステム中に負荷される。これらの値は後でアクセスするために1組のシステムバッファのような第3のメモリ位置中に蓄積され、読み出される。システム動作パラメータは、中央プロセッサアクセス時間の予め選択された最小値および最大値、プロセス内部セット中のメモリの頁の最小数、内部セットインクレメントの大きさ、内部セットデクレメントの大きさ、最大および最小頁事故割合等のための制御値を設定する。特性パラメータは、頁事故割合、頁事故待機、内部セットの大きさ、休止待機、計算可能なプロ

において実現される。リソースはCPU処理サイクル、動作セットメモリの大きさおよび処理時間、プロセス優先度、およびI/O転送タイミングを含み、そのシステムリソースの割当てに関して各プロセスに対する一連の内部特性値を測定するために1以上の監視ルーチンを有する。予め定められたシステム制御パラメータの調整によつてシステムリソースをダイナミックに再割当てする方法は、メモリ位置に所望の特性値のセットを蓄積し、第2のメモリ位置に各プロセスに対する現在のシステム制御パラメータに応じて測定された監視された特性値を累積し、対応する測定された特性値と所望の特性値の間の予め定められた関係にしたがつて調整係数を発生し、この調整係数にตอบสนองしてシステム動作パラメータを調整するステップを含む。

好ましい実施態様においては、この発明の方法はさらに予め定められた時間待機して前記累積ステップに戻り、前記待機の期間は前記所望の特性値と測定された特性値との相対的な差に応じて複

セス、計算可能なプロセス外部交換状態、頁衝突、および各プロセスに対する自由頁に対するパラメータを含んでいる。

前記調整ステップは、さらに割当てられたCPU時間と関連する高低限界を越える頁事故割合のプロセスに対して動作セットの大きさおよび量子およびI/O T Aパラメータ値を調整し、現在のCPU時間要求の観点における所望の頁事故範囲限界内のプロセスに対して動作セットの大きさおよびアクセス時間を減少させるステップを含んでいる。動作セットリストに掲載された過剰なメモリは自由頁使用のために転送される。

この発明の方法を実行するための装置は、中央処理装置(CPU)と、このCPUに接続された頁アドレスされる実上のメモリと、システムリソースへのアクセスを割当てるための予め割当てられた制御値を有するシステム動作パラメータの固定されたセットとを有し、リソースはCPU処理サイクルおよび内部メモリを備え、システムリソースの割当てに関する各プロセスに対する1組の

測定された特性パラメータに対する値を発生するために1以上の特性モニタを使用して多使用者コンピュータシステムにおけるリソースの割当てを制御する。この装置はCPUに接続されて動作パラメータおよび対応するシステム制御値を蓄積するシステム蓄積手段を使用する。CPUに接続されている状態蓄積手段が各プロセスに対する特性パラメータに関連する値を蓄積する。CPUに結合されて動作するモニタ手段が各アクティブプロセスに対する前記特性値を測定し、それらを状態蓄積手段中に蓄積する。状態蓄積手段はモニタ手段に接続されて測定された特性値を受信し、1組のシステム状態値を形成するためにそれらを累積することができる。調整手段はCPUとプロセスとの間に接続され、プログラムの制御下にプロセス制御パラメータを調整するために動作システムパラメータを使用する。CPUおよび調整手段に接続されているリソース制御手段は各プロセスに対するリソース割当てを調整するために両者に対する所望の値の新しいセットを与える。応答手段は

周期的に特性値を累積し、プロセスのリソース割当ての変更において調整手段により使用されるための調整係数を発生する。

この発明のさらに別の観点においては、装置はまた調整手段およびモニタ手段に接続されて前の調整値に応じてモニタおよび調整の周波数を決定するモード選択手段と、情報を蓄積するための少なくとも1個のバッファとを備えている。

【実施例】

この発明はデータおよび相互作用特性のスループットを改善するために多使用者コンピュータシステム中の計算リソースをダイナミックに再割当てする方法を提供する。この発明は、予め設定された特性基準からの現在のプロセス特性の偏差を測定する緊密に結合されたフィードバックループと共同して動作する各アクティブプロセスに対して動作セット調整ルーチンを呼び出すことによつて行われる。この方法で、この発明の方法は歴史的なシステムの使用に対するメカニズムを提供し、システム動作パラメータ値をダイナミックに調整

するための基礎として特性情報を処理し、システム要求に応じてそれらの値の特別の実行を処理し、システム全体の性能および応答特性を改善する。コンピュータシステムは動作状態に向かって調整され、それにおいては周期的に測定された特性パラメータ値が所望の1組の値に等しく、または近似される。測定された特性が予め設定された基準に近似したとき、プロセスはもはやシステム負荷が特性に顕著な変化を生じるか、新しい所望の特性値のセットが与えられるまでは調整されない。

この発明の方法は実質メモリ型のコンピュータシステムを参照して以下説明される。この発明が使用されるコンピュータシステムの一例は、VMS動作システムを使用するデジタルエキップメント社のVAXファミリーのコンピュータである。明瞭にするために構成の詳細を明らかにし、コンピュータシステム動作における改善を示すために、この発明はVAX/VMSバージョン4.xx環境で実施されるものとして説明されるが、当業者はこの発明の方法が容易に他の多使用者コンピ

ュータ動作システムにも同様に適用可能であることが認められるであろう。

これらのコンピュータシステムは多数の使用者および/またはプロセスを有してそれらは任意の時間に制御され、または動作される。プロセスという用語は、プログラムが特定のターミナルへ相互作用なしに負荷されおよび実行されることがするためにあるターミナルにおける末端使用者以上の存在を示すために使用される。また中央プロセッサに接続され、動作する多数のプリンタ、モデム、および磁気媒体蓄積装置がある。これらはCPUが相互作用するプロセスまたはプロセス制御の下でそれぞれ走行する。

コンピュータ技術でよく知られているように、実質メモリまたはメモリコンピュータシステムはそれが内蔵されているアクティブなメモリか、ディスク駆動装置のような遠隔蓄積装置上の2次メモリであるかに関係なくあらゆる利用できるメモリ蓄積位置に処理メモリアドレス割当てによつて区別される。そのようなアドレスの全てのリストは

コンピュータシステム上のプロセスによりアドレス可能な“実質”メモリの装束を形成し、1以上の変換表またはリストは実質メモリアドレスを実際の物理的位置へ変換するのに使用される。

この方法において、システムで実行されるプロセスは全てのメモリを連続するものとして扱い、任意の時間に実際に内蔵されている量よりも大きい量のメモリをアドレスすることを許容する。処理のために内蔵されていないメモリから内部メモリへ情報を転送しなければならないことによつて生じる遅延は通常は低く、一連の特定された待機状態によつて容易に適合される。

動作および制御を容易にするために、この形式の動作環境における全てのメモリは一定数のデータバイトからなる“頁”に分割される。頁データは特定の位置にある必要のない、予め定められたバイト数に対する単なる一連の連続したアドレス可能な位置である。現在のVAX/VMS型動作システムを使用するコンピュータは典型的には512バイト程度のデータを含む頁を使用し、一方

他の動作システムは典型的には1024乃至2048バイトの程度の頁を使用するが、その他の大きさも可能である。大きな頁はメモリの集中的処理のために内部メモリに入出力する多量のデータを保持しシフトするには非常によく適合するけれども、少量のデータが多数のプロセスにより使用されている場合には潜在的にメモリを無駄にする。

メモリ頁の管理において、各プロセスはデータを蓄積し動作を行う内部メモリ中に1組の頁が与えられる。実行プロセスはコンピュータシステムが内部メモリから割当てることができるデータの蓄積および計算の遂行のための頁の決定可能な数を要求する。しかしながら、内部メモリの量は限定されているから、任意のプロセスで使用される頁の全体数の一部は内部メモリ中の一時に多数の処理に適合するように所定の時間に遠隔蓄積または非内蔵メモリにある。所定のプロセスに割当てられた内蔵頁は頁の“内部セット”または“動作(working)セット”(WS)として知られている。

プロセスがアクティブにされ、すなわちメモリ中に負荷されるとき、自由頁リスト(FPL)の一連の内部頁はそのプロセスに割当てられ、予め定められたデータ頁数が内部メモリに負荷される。割当てられた頁に対する一連の位置またはメモリアドレスポインター処理のために動作セットリスト(WSL)中に蓄積される。プロセスはこれらのポインターを使用して必要に応じてデータを割当てられた頁中へ転送する。割当てられた頁の全てが実際の処理の要求に応じて内部セット中で使用される必要はない。WSL上にリストされた頁は依然としてプロセスにより使用されるまでは自由メモリの一部である。

プロセスがデータ蓄積または検索のために内部頁限界を越えるとき、すなわち全てのWSL頁が使用されるときまたは内部でないデータに対して呼ばれるとき頁事故が発生し、データは動作セットと内蔵されていないメモリとから部との間で交換されなければならない。この方法は一般に良好に動作する。何故ならばプロセスは任意の1時期

にそれらの頁の全てを要求することはなく、しばしばI/O転送、他の処理を待ち、或いは周期的に待機状態にされ、そのため頁交換によりわずかな遅延が生じる。

しかしながら、比較的小さい動作セット、低い頁数を有するコンピュータシステムにおいて多数のプロセスが存在するとき、CPUは内部メモリと非内部メモリとの間のデータのI/O転送の管理に費やされる時間量が増加する。この形式の管理のためのCPUサイクルの損失は全体のシステム応答性に作用する他のシステムに対して利用されるCPUリソースの量、CPUサイクルの数を減少させる。また内部メモリとの間の頁の交換は磁気媒体における読取り/書込み動作を必要とし、それはまた特定の処理およびI/O管理用の一般的処理の両者に対するシステム応答性を遅くする。モニタルーチンまたは装置は1以上の蓄積レジスタまたはメモリ位置中の頁事故データを累積するように動作システムによつて使用するために設けられ、そのため頁事故の数およびWSメモリの

大きさおよびCPUアクセス時間の関数としてモニタされることができる。これは処理効率および適切なWSの大きさの指示を与える。この情報はコンピュータシステムにおいて動作セットの大きさの調整に使用されることができ、高い負荷を緩和し、処理遅延を最小にする。

それ故メモリリソースを制御または管理するために、動作システムはシステム制御パラメータを設定し、その値は動作セットの大きさを決定する。動作システムパラメータはまた、動作セットが変更され、増加または減少される頁数の数およびそのような動作のタイミングを制御する。同時に、動作システムは時間の最小および最大量を制御することができ、プロセスは動作セット中で付勢されることによつて計算または実行可能である。

これらのパラメータはコンピュータ動作システムが設備され、平均または典型的なコンピュータシステム要求に一致される時に設定される。これらのパラメータはポインターの表またはリストによつて対応する制御または動作値が蓄積される特

定のメモリ位置またはレジスタとリンクされる。パラメータに対する値はコンピュータシステムが電力設定またはリセットのような初期化される都度当業者によく知られている方法によつて初期化される。パラメータ値の表は典型的には核心理心(kernel)または実行(exec)と呼ばれるコンピュータ動作システムの階級の最高次モードの部分として処理され、使用者によつて変更されたり、外部者によつて妨害されることを防止するために符号化された限定されたアクセスを有する。しかしながら、プロセスが“権限のある(オーソライズされた)”場合またはシステムが動作システムの核心理心モデルレベルにアクセスする場合、これらのパラメータはそのプロセスによつて処理されることができる。

前述のように、コンピュータシステムはプログラムまたは動作環境を初期化することによつて、および動作システムを発生または負荷することによつて動作を開始する。すなわち、それは予め定められたセットの命令、符号および関連する動作

パラメータに対する制御値をリザーブされたメモリ位置へ負荷し、それからそれらはプロセスおよびリソース割当てを制御するために使用される。多数の固定パラメータが基本CPUサイクルタイミングまたはアクセス、全体のメモリの大きさ、またはI/Oしていおよびタイミングのような機能を制御する。基本タイミンググループ、割当て、または物理的割当てはハードウェアによつて固定され、コンピュータシステムに対して各種のI/Oルーチンを適切に使用するため、および使用者間で調和されなければならない状態を待つためにソフトウェア中で固定されたままでなければならない。しかしながら、初期化において蓄積されたいくつかの調整パラメータがある。これらのパラメータはコンピュータシステムの同期を助けるためにシステム管理者によつて変更可能である。VAX/VMSコンピュータシステムでは、これらの調整可能なシステムパラメータのいくつかは量子、IOTA、PFRATL、PFRATH、WSDEC、WSINC、AWSTIME、AW

SMIN、BORROWLIM、GROWLIM、MPW-THRESH、およびPIXSCANとして示されている。

量子パラメータは典型的に時間の最小量を定め、処理のためにCPUにアクセスすることを許容する。現在のVAX/VMS構成においては、このパラメータは、それぞれ10ミリ秒の単位で内部メモリ中の非待機状態のプロセスが付勢される時間の最小量を設定する。しかしながら、他の時間間隔は他のシステム設計に使用されることができる。IOTAパラメータは量子時間に対するチャージとして作用し、プロセス実行または付勢状態に対して待機状態への周期的エントリーを確実にするための外部時間限界を設定し、そのため各プロセスは優先度および活性度に関係なくCPU時間を規則的に自由に使用する。このパラメータはまた各10ミリ秒の単位で期間を選択する。量子パラメータに割当てられた値によつて特定された期間が一度経過されたならば、プロセスはエンド量子フラグまたは指示を設定しCPU制御を他のプロセ

スに送る。

WS短縮を伴う上記リストのパラメータはアクチブプロセス動作セットに対する動作限定を設定する。AWSMINパラメータに対する値は動作セットに対する最小の大きさを設定し、それは自由頁リストFPLからWSLへ割当てられたシステムにおける内部頁の最小数である。WSDECおよびWSINCパラメータに対する値は、WSまたはWSLにおける頁に対するポインターが動作セットの大きさが調整されるときには常にそれぞれ減少または増加される頁数を待定し、AWSTIMEは10ミリ秒の期間で動作セットの大きさを調整する間の最小待機時間を設定する。

上記リストされた残りのパラメータは、WSの大きさ、調整、および頁アクセスに関係する他の限界を決定する。PFRATLおよびPFRATHパラメータ値は10ミリ秒当りの頁事故の数に対する低および高限界をそれぞれ設定し、それ以下および以上に動作セットのための調整が制定される。

御装置16中にWS制御値を蓄積し、単一プロセスに対する一連の蓄積レジスタ18中のプロセス状態に情報を累積する。プロセス制御ブロック14はまたライン22によつて特性モニタ20に接続されて蓄積レジスタ18から情報を受ける。

コンピュータシステムプロセス制御装置10およびプロセス制御ブロック14は一般にソフトウェアルーチン、メモリ位置、およびシステム動作を制御するため動作システムによつて使用されるハードウェア装置を備えている。これらの制御の使用および設計は当業者によく知られており、ここではさらに詳しい説明はしない。

動作システムパラメータ制御装置12は固定された動作システムパラメータPFN12aおよびダイナミック動作システムパラメータPDN12bのセットまたはアレイを含む。コンピュータシステムの初期化において、これらの動作システムパラメータコンピュータ動作システムOSのソフトウェアまたはファームウェアにより前記のメモリ位置中へ負荷される。パラメータは一般に動作システ

GROWLIMパラメータ値は、ページシデン
ト (pagesident) の最小数SUSPを設定し、または浮遊および外部交換SUSPOである。プロセスはまた処理において長い遅延を有することができ、動作システムがそれを“休眠”および外部交換モードHIBOに置くことを要求し、或いは外部交換LEFOを行わせる局所的な事象に遭遇する。勿論動作システムは、この発明のこの構成では説明しない種々のプロセス待機状態を構成するためのMWAITまたはMPWのような付加的な待機状態を維持する。

リソース割当てを制御するための多使用者コンピュータシステム中の制御パラメータおよび待機状態と関連したこの発明の動作は、第1図のブロック図を参照することによつてよりよく理解されるであろう。第1図において、コンピュータシステムプロセス制御装置10は、一連のプロセス制御ブロック(PCB)14と並列に接続された動作システムパラメータ制御手段または制御装置12を使用している。各プロセス制御ブロック14はWS制

ムソフトウェアの初期化の一部として負荷されるが、当業者にはそれらがPROMのような各種のハードウェアメモリ装置中に予め蓄積されて自動的に登録またはアクセスされることができることが理解できるであろう。

ダイナミック動作システムパラメータPDNはシステムの広い範囲の動作または機能を制御する値を有する調整可能なパラメータを有する。ダイナミックパラメータの正確な数は使用される特定の動作システムに応じて変化し100以上程度に大きくできる。好ましい実施例では、前に説明した中のただ10のパラメータがこの発明の方法の実行において調整されるに過ぎない。しかしながら、コンピュータシステムは発展し、速度、メモリ容量、設計の効率、新しいパラメータまたはパラメータラベル等を含む動作システム構成における変化を受けるので、他のパラメータがこの発明による調整のために選択されることができると。

動作システムパラメータは各プロセス制御ブロック14によつて使用されてそのコンピュータシ

テムでアクティブである各プロセスに対する動作限界または特性を設定する。これは初期化されたときに各プロセスに割当てられたメモリ位置またはレジスタ18のセットにおける動作設定制御値を設定し、蓄積することによって行われる。これらの制御値は最小の動作セットの大きさおよび動作システムパラメータにより決定されるCPUアクセス時間関係する制御限界を含む。しかしながら、WS制御値は頁事故を減少させるように動作セット頁リストを増加させるために、または可能であるメモリ再割当てのための動作セット頁リストを減少させるために、歴史的プロセス特性に応じて変更される。このためにプロセス制御ブロック14はプロセス調整装置(PA)24を介して動作制御システム手段12に接続され、このプロセス調整装置24はWSDECおよびWSINCのようなダイナミックパラメータ値を使用して適切に動作セットの大きさを調整する。動作システムパラメータは一般にライン28によつてプロセス調整装置24に与えられるが、初期化のため、または直接制御が

所望される場合にはライン28で直接供給されることができ。

各制御ブロック14はまたプロセス状態蓄積レジスタ18中に情報を累積する。累積されたデータは固定された時間間隔にわたる頁事故の数、現在の待機状態等に関する情報を表わす。状態蓄積レジスタ18はライン22によつて特性モニタ20に接続され、また以下説明するように累算器36に接続されている出力を備えている。

特性モニタ20は頁事故(PF)、サイクル、プロセス頁カウント(PPGCNT)、グローバル頁カウント(GPGCNT)WSLの大きさ、自由メモリ(FREE-MEM)、自由頁リスト(FPL)の大きさ等のようなパラメータの表を維持し、それらは各プロセスNの特性を決定するために使用される。もしもプロセスNの頁事故の割合が動作システムに対する初期値によつて設定されたときに高くなり過ぎれば、WSDECおよびWSINCは動作システムパラメータ制御装置12またはプロセス調整装置24によつて使用され

ることができ、動作セットの大きさを調整するためにWSレジスタ18中に制御値を再び蓄込む。

説明した動作システム管理方法はプロセスにわたるメモリ再割当てのためにプロセス動作セットの大きさの調整を可能にするが、多数のプロセスのためのダイナミック負荷の問題を解決するためには歩適当であることが認められた。上述のように固定された動作システム調整パラメータは全ての調整されたプロセスに均等に“平均解決”調整を供給し、プロセス要求における変化或いはダイナミックに変化する負荷を考慮していない。WSINCに負荷された値は比較的大きく、WSDECは非常に小さく、そのためアクセスされない頁に対して、非メモリ集中処理に対してわずかの追加頁が必要とされるときに多数の頁ポインターが不必要にWSLに加えられる。

プロセス調整はまたプロセスがエンド量子状態に到達するまで許容または達成されない。これはそれらの量子限界に到達しない浮遊または休眠状態の多数のプロセスが調整されないことを意味す

る。これらのプロセスはしばしばリソースを結合し調整を必要とするプロセスである。それ故現在の動作システム限定はコンピュータシステムは負荷不均衡の多数のソースのアドレスをできなくしている。この発明はエンド量子状態に対してのみ調整することに限定されない。

さらにWSINCとWSDECとの間の不均衡は振動を生じさせることになり易い。即ち、過剰な頁事故を避けるために必要な最小の頁量より少し下に動作セットが減少したとき多数の頁が動作システムによつて加えられる。必要でない頁を除去するためにいくつかのWSDECの供給が行われ、一般に動作システムは再び最良の動作セットの大きさをオーバーシュートする。これは多過ぎる頁と少過ぎる頁との間の一定の振動になり、大量のCPU時間とメモリリソースを消費し、一方WSLは一定して変化される。それ故デジタルエキップメント社のようなコンピュータ製造業者はどんな形式でもダイナミック調整を試みることを推奨している。それはシステムに一定してさらに

要求を生じさせ、負荷の問題を増加させるからである。その代わりにダイナミック動作システムパラメータ値はシステム管理者によつて長期間に亘り調整され、ダイナミックシステム負荷または要求にたいする価値に欠けている。

この発明は、WS動作限界を制御するために使用されるダイナミック動作システムパラメータ値を自動的に調整するために各プロセスにおける特性情報および統計コンピュータシステム情報を使用することによつて現在の固定した調整方法によつてこれらおよびその他の問題を解決するものである。調整はコンピュータリソースを再割当てするために短期ベースで行われる。個々のプロセスベースにおけるフィードバック制御された再割当ての使用は簡単な動作システムを越えたリソース管理を許容し、平均自由頁リストの大きさを増加させ、処理遅延を減少させる。

これはダイナミック負荷平衡装置(DLB)30を設けることによつて行われ、それは動作システム制御装置12、プロセス調整装置24、および特性

作システムの最高のモデルレベル、核心、または主要部に変化を行わせるために高い優先度およびアクセス権限を備えている。この権限なしに調整パラメータ変化は許容されない。

DLB30は割当てられた蓄積の8000バイト程度で走行する小さな分離したプロセスに対してのみ充分なメモリを取る。DLB30は動作システム制御装置12を完全に置換することなく動作システムそれ自身を要えることもないことが重要であることを注意すべきである。これは他のプログラムおよび設備が計画された、または予期されたものとして動作することを確実にする。システム事故を生成し、一般的な機能または保守契約を無効にする動作システム変更は行われない。

この発明の方法は第2図にフローチャートの形態で示されている。この発明の方法にしたがつてコンピュータシステムを動作させるために必要な機械実行可能な命令を過度の実験なしに当業者が処理できるように、一連の疑似コード表示が明瞭に説明のために付属する説明と共に表IおよびV

モニタ20に結合され、それらと相互作用して動作セット制御値を調整するために使用されるパラメータを調整する。ダイナミック負荷平衡装置30はまたプロセス状態レジスタ18からデータを受けるために接続されてコンピュータシステムについての統計的動作データを累積する。

定められシステム中に負荷された初期コンピュータシステム動作パラメータにより、この発明は予め選択されたプロセス名を使用して呼出され、または付勢される分離したプロセスとして構成される。プロセスは初期化ルーチンの一部として自動的に、それがシステム管理によつて所望されるときに常に呼出されることが出来る。分離したプロセスとしての動作はまた所望されたときこの発明の方法の順序正しい撤回または消勢を与える。完全に独立したプロセスはターミナルI/O限界、モニタ、またはオーバーヘッドにさらされることがなく、その自動なダイナミックな調整動作中CPU時間を消費することが少ない。

しかしながら、この発明の方法および装置は動

に示されている。この発明はいくつかの異なるコンピュータ動作システムにおける各種言語で構成することができ、単一ソースおよび/または動作構成の目的コードによるこの発明の説明は非常に詳細な事項によつてこの発明の主題を不明瞭にしこの発明の理解に必要な明瞭性および簡潔さを損うことになるから、特定の目的またはソースコードは示されていない。この発明の方法は重要な多使用者コンピュータシステムに利用できるコンピュータ言語および命令を使用して実行される。好ましい実施例はVAX/VMS動作環境で使われるBLISS言語で行われる。

表Iは付勢または再付勢されたときに使用されて必要なバッファメモリ位置を設定し、その方法に使用された各種のパラメータ値を負荷かし検索する初期化ステップの例が示されている。

表 I

初期化

初期化：過程

蓄積初期化、

新旧システム動作パラメータ値蓄積
新しい値をシステム制御装置へ提供
実行

バッファ初期化

アドレスチェック、バッファ未発見のとき
ストリング認識、その後新バッファ初期化
終了

その他

バッファアドレスと認識ストリング蓄積
新動作システムパラメータ値蓄積
新DLBプロセスパラメータ値蓄積
旧動作システムパラメータ値取得

終了

DLBプロセス用のプロセス情報取得

システム情報取得

MPW高限界取得

MPW低限界取得

バージョン取得

BALSETCNT取得

CPU取得

終了

DLB状態チェックが誤りなら、または
動作システムが悪状態なら、

退出

その他

BALSETCNT取得、ADJなし、ADJなし

ADJなし、およびADJベースなし

設定: ADJ-MEMなし

(BALSETCNT×ADJ-INCRなし) +

ADJベースなし

設定: 最後のDALSET←BALSETCNT

設定: 最後のABANDONED

設定: 新対応動作システムパラメータ値

終了

割当てられた不使用メモリのDLBプロセス
WSへの除去

終了

第2図および表Iに示すようにDLB30は分離
プロセスとしてまずステップ40で付勢され、それ

からステップ42でコンピュータシステム上のバッ
ファ32の存在についてチェックされる。これはア
ドレス値を渡すことによつて行われ、予め定めら
れた論理ラベル、この場合DLB-SYSBUF
-ADDRしたの認識ストリングはDLB30の
スタートにおいて却下される。これはDLB30が前
に走行していたとき生成されコンピュータシス
テムにおいて依然として付勢されている可能性のあ
るバッファを利用するためになされる。バッファ
は浪費されてはならないコンピュータシステムの
ための価値あるリソースまたはメモリアーバ
ードを喪失す。これは、この発明の方法が終了ま
たは消勢され、バッファ32に対して使用されたメ
モリ位置またはアドレスが再付勢の前にコンピ
ュータシステムによつてまだコリアされていない場
合に生じる。

バッファが発見されなければ、DLB30による
データまたはパラメータ情報の蓄積のための一連
のメモリ蓄積位置32a, 32b, および32cを有す
るバッファ32がステップ44で初期化される。これ

はWSに対して500頁程度のメモリを要求する
内部のみのプロセスとしてステップを処理する
DLB30を負荷することによつて行われる。これ
らの頁のいくつかはこの発明の方法の実行のため
にパラメータ値およびシステム情報その他の蓄積
および検索のために使用される。

初期化の一部として、コンピュータメモリ位置
は論理的“DLB-SYSBUF-ADDR”ス
テップ他で指定され、2個の連続するデータワ
ードを蓄積するために使用される。第1のワードは
バッファ後とであり、第2のワードはダミーまた
は認識値である。ダミー値は都合のよい選択を表
わす123456789を有する選択された蓄積の大きさ
内に適合する任意の文字のセットから構成される。
この文字ストリングは負荷平衡プロセスが付勢さ
れたとき不正確なアドレス値を使用することを阻
止するために、バッファの存在をテストするため
に使用される。

バッファメモリ32aは新しく初期化されたもの
でも古いものでも以下の表IIに示す予め定められ

た動作システムパラメータ値のセットでステップ48において負荷される。動作システム制御パラメータいくつかの典型的なコンピュータシステム欠陥値と共にバッファ32a 中蓄積された新しい値を示している。

表 II

パラメータ	新しい値	欠陥値
量子	2	10 (10ミリ秒)
IOTOA	1	
PFRATL	0	0事故/10秒
PFRATH	10	150事故/10秒
VSINC	305	150頁
VSDEC	0	35頁
AVNHIN	10	50頁
AVSTINE	5	20
PIXSCAN	baisetent	
MPW-THRESH	18000	200頁

表IIの動作システムパラメータは動作システム制御装置12中で発見された対応する動作システム

プロセス調整間の待機は200 ミリ秒から70ミリ秒 (AWSTIME) に欠陥値が減少してダイナミックな負荷変化に迅速に反応することを可能にする。プロセスがCPU時間を割当てられた時間量は量子パラメータにおいて減少する。これはコンピュータシステムが短期間の処理からシステムリソースにおける負荷かまたは負担を迅速に除去することを許容する。

MPW-THRESHのような他の動作システムパラメータ効果メモリ割当ては内在しないメモリに対する頁の溢れて続いて自由頁リストに転送されるのを減少するためにより高い値を割当てられる。これはソフト頁事故に利益となるようにハード頁事故における頁を交換するために必要なCPU時間を減少させる。PFRATLおよびPFRATHパラメータ値はさらに頻繁な間隔で動作セットの大きさの調整を許容するように調整される。

動作システムパラメータに加えて、計算またはDLBプロセスパラメータと関連する値の他のセ

パラメータを使用される。これらの値は特定のコンピュータシステムまたはあるクラスのそのようなコンピュータシステムのいずれかに対する一般的動作特性の注意深い考察後に到来する。好ましい実施例では、開示された値のセットはVAX/VMS型動作システムに適合でき、多数の子となるコンピュータシステムに対して改善された動作を行うことが発見された。

バッファ32に蓄積された初期化値はリソースのより良好な割当てのためにシステムパラメータを変える別のステップで使用される。このラインに沿って動作セット調整パラメータWSDECおよびWSINCに通常割当てられる値は動作システム調整ルーチンをディスエーブルにするように調整され (WSDEC=0)、調整を行うべきより大きな頁数 (WSINC=503) に対する容量を与える。最小のWSの大きさは典型的な50頁の欠陥値から10頁に減少する。これは多数の不使用頁を示すWSLを有することから無メモリの集中的処理を阻止する。頁は必要により加えられる。

ットはバッファメモリ32b に負荷されてDLB30によつて使用され、これらは末尾の表IIに示される。これらのパラメータはこの発明の調整およびタイミングにおける限界を設定するために使用される。

これらのパラメータは、心配な (ウォーリー) モードの待機のためのウォーリー待機、パニックモード待機のためのパニック待機、および以下さらに説明するような使用する適当な待機モードパニックMEM、ウォーリーMEMを決定するメモリ量のようなある種の待機期間の長さのような機能を制御する。WS頁調整、ADJ-INCRなし、およびある境界内にメモリ調整プロセスを維持するために使用されるADJベースなしのベース量に対して2個のパラメータ制御またはセットアノーマルインクレメントが使用され、一万金でのプロセスは調整され、新しいシステム情報は特に非常に大きい自由頁リストを有するシステムに集められる。ベース、最小、および最大メモリ調整量MEM-ADJ、MAX-MEM-ADJ、

およびMIN-MEM-ADJおよびメモリのための最小値、絶対メモリ、メモリ調整、およびCPU時間の大きさMEM-MIN、ABS-MIN、ADJ-MIN、CPU-MINのための一連のパラメータが値を設定される。最小および最大優先度および事故割合、MIN-PRIOR、さMAX-PRIOR、MAX-FAULTSもまた指定される。グループパラメータはシステムプロセスと考えられるプロセスと残りの使用者プロセスとの区別を許容する。現在の実施例において、システムプロセスは調整されず、このパラメータの認証はそのようなグループに分割されるときプロセス間の文化を許容する。2つの非常に重要なパラメータ値がそれらのDLB-WSDEC RおよびDLB-WSINC Rパラメータに割当てられ、それらは動作システム制御パラメータとは無関係にDLB 30によりWS調整のためのインクレメントおよびデクレメントを設定するために使用される。

第2図では次のステップ50ダイナミック動作パ

メモリから検索される。これは動作セット情報のみならずプロセスの現在のモード、核心レベルまたは規則的OS、および待機状態に関する情報の収集を含み、処理の変更を阻止し、一方ではパラメータは調整される。

n個のプロセスのそれぞれに対する状態か情報はまたアキュムレータ36に累積され、パラメータ値を調整するために統計的に発展するように特性値の発生のプロセスを開始させる。

利用できる自由頁、全体的システムメモリ、および一般的システムおよびメモリ割当てのような他の基本的システム情報もまたチェックされてもよい。所望の場合にはこの発明の方法は認証ステップを使用して遅延数またはその他のDLB 30に対するソフトウェア識別を確認し、コンピュータ動作システムによる獨立性を確実にする。何故ならば、パラメータラベルおよび値はシステムによって異なるからである。

この発明のプロセスは1以上のシステム動作パラメータをダイナミックに変更または変化させる。

ラメータに対して動作システムにより現在値セットを検査し認証する。これらの値はバッファメモリ32c中にコピーされ或いは必要に応じて後でアクセスできるメモリ位置またはアレイ中に蓄積される。この情報は累積され、そのためコンピュータシステムは通常の方法でもとの状態または所望のときこの方法の終了においてパラメータ値のセットに戻ることができる。同時に何等かの変化がなされる前に所望のパラメータ値のセットと比較されることを許容する。初期動作システムパラメータの長期項遠隔蓄積値は必要ではない。何故ならばそれらはシステムが再よき化され、または何等かの理由でリセットならば初期状態に自動的に戻るからである。

モニタ20により各プロセスのための特性パラメータに対して割当てられた値はまたステップ50でバッファメモリ32c中にコピーされる。各仕事またはプロセスの現在の状態はチェックされ、プロセス調整中にメモリ中に記録され、必要なときにモニタ20により蓄積されたコンピュータシステム

調整されたパラメータは、リソース割当てに関しコンピュータの動作に最も容易に、または著しく影響する動作に基づいて選択される。

好ましい実施例は、DLB 30が動作できる適当な部屋を与えるためにDLB 30によつて調整されない動作システムパラメータのいくつかに予備調整を行う利点を有する。すなわち、最大の動作セットの大きさまたはダイナミックメモリ割当ては調整されてDLB 30に変化を行わせるために適当な部屋を与える。もしもこのようなシステムパラメータが小さ過ぎる値を割当てられるならば、DLB 30は効率的にリソースを再割当てするために他のパラメータに対して変化を行わせるように厳しく制限される。

これはダイナミックメモリに対する処理要求の速度を増加するために通常使用されるメモリパケット行になつたリストに対する値を含んでいる。これらのパラメータは小さい要求パケットリスト、SRPCOUNT、SRPCOUNTV、I/O要求パケットリスト、IRPCOUNT、IRP

COUNTV、および大きな要求バケットリスト、LRPCOUNT、LRPCOUNTVを指定するために使用される。これらのパラメータに対する値は頁化されていないメモリスペースを割当てるよりもコンピュータ中のリンクされたリスト上の多数のバケットを確保することによつてメモリアクセスを改善するように調整される。頁化され、および頁化されていないダイナミックメモリ最大WS、実質頁メモリおよび表IVに示された関係するパラメータもまた変化される。表IVに示された値は好ましい実施例を最も有効にする最小のガイドラインである。

第2図に示す次のステップ52は、ダイナミック頃も動作パラメータを重ね書きし、ステップ48に示されたようなバッファ32中に蓄積された新しい動作または制御値にそれらを設定することである。場合によつては新旧パラメータは前のシステム同型によつて同意されているが、これらの値はさらにこの発明がシステム動作特性に影響するパラメータを使用している特定のダイナミックな方法の

自由MEMを含む。もしもDLB39が自由メモリの値がN-ADJ-MEMパラメータに対して設定された値よりも大きいことを発見したならば、この時点では調整は行われず、DLB30は待機モードに入る。

他方DLB30は最終平衡セットカウントLAST-BALSETと排気されたプロセスの数に対する最後のカウンタLAST-ABANDONEDとの間の差に対してアクティブカウント値をリセットするように前進する。しかしながら、1の最小値は強制される。

安全メモリ量MEM-SAFE-AMTはプロセス呼びのため留保されるべきメモリの量に対するバッファ値を与えるようにセットされる。表Vから明らかなように、MEM-SAFE-AMTパラメータに対する値は最小メモリ値MEM-MINプラスアクティブプロセスの数の70%で割算された自由メモリ値(1より小さくはない)に等しく設定される。これはアクティブプロセスの約70%までが次の過程においてメモリを要求するとい

ために一般的に相違していることが予想される。

この点においてDLB30はそれ自身のWSL中の頁をチェックし、できるだけコンパクトな動作エンティティを与えるために必要ないものを除去する。DLB30はコンピュータシステムのリソースの改善ができるだけ小さいように使用されるように設計される。この使用しないメモリは通常のコンピュータシステムのページルーチンを使用してステップ54で除去(ページ)されコンピュータシステム自由頁リストに戻る。個々のプロセス調整のためにこの発明で使用されたステップに対する疑似コードリストは末尾の表Vに示されている。

表Vおよび第2図から明らかなように、DLB30はステップ56においてサイクルカウンタをインクレメントし、動作システムまたはモニタ20からのシステム情報を検索することによつて調整プロセスを開始する。この情報はシステムにおけるアクティブプロセスの数、アクティブカウント、頃もにおける計算可能なプロセスの数、COMカウント、および全体および自由メモリの両TOT-MEM、

う事実に基づいて安全メモリを算定している。DLBは迅速なダイナミック端か変化に対して調整するためにその割当て計画で十分な自由メモリを維持する。

調整されたプロセスの迅速な再スケジュールを助長するためにさらに処理する前に、ステップ58で値子およびIOTAパラメータに対して新しい値が設定される。

DLB30はステップ60で表Vに示された疑似コードのように個々のプロセス調整の実行に進む。検索されたパラメータ値のいくつかはプロセスid、PID、プロセスCPU時間、CPUTIME、PAGEFAULT、プロセス頁カウント、PPGCNT、グローバルな頁カウント、GPGCNT、および優先度ベース、PRIBを含む。

ステップ60では、第1のプロセスによつてスタートし、それらが全てチェックされ調整されるまで連続するので、DLB30はコンピュータの動作特性およびモニタ20とプロセス状態レジスタ18により蓄積された値を読み取ることによつてコンビ

ュークにおける各プロセスの動作特性をチェックする。

交換されたか否かを決定するために追加的な情報があるめる前に各プロセスの状態をチェックするか、またはさらに情報がそのプロセスで検索される前に中断されることのみが重要である。これは頁カウント、頁事故等のようなプロセス情報についての質問の作用がアクティブでないプロセスと共に多くの場合そのプロセスを再活性化させるという事実による。これはこの発明の目的が不必要なリソースの使用を減少させることであるから自滅性である。

一度プロセスの状態が知られたならば、依然として活性であり待機状態にないこれらのプロセスは現在の状態がどうであるか、特性がどうであるかを見るためにチェックされる。各プロセスの特性パラメータ値は所望の特性基準のセットと比較される。

プロセスWSの現在のメモリが絶対最小WSの大きさに対するパラメータに等しいか小さい場合

プログラムはコンピュータシステムWS調整ルーチンである。VAX/VMSにおいてこのルーチンはsysadjustと呼ばれている。しかしながら、他のコンピュータシステムは同様のルーチンを使用しておりよく知られている。

DLB30は表VIに記載された基本プロセスを検索する。この表VIは最後に調整されるからプロセスの特性を示している。この情報はDLB30によって蓄積されて後で次の過程のものと比較される。同時に前の過程でこのプロセスに対して蓄積された情報は頁事故、CPU時間、および頁カウントのようなある種のパラメータの比較のために検索される。

最後の時にDLB30は調整を行っているから、発生した頁事故の数およびCPU-AMNT中の変化に基づいてMEMAMTパラメータに対して選択された値が決定され、したがってADJSMAMTパラメータが決定され。この後者のパラメータの値はプロセスID、コンピュータ上のインデックス位置、DLB-WSINCRと共に調整

にはさらに変化は行われない。もしも現在の頁事故の数が前に測定されたものよりも大きく、CPU時間の現在の量が最小CPU時間の値に等しいか、それよりも高く設定されるならば、変化は行われない。プロセスに対する現在の優先度がMIN-PRIORおよびMAX-PRIORによつて設定された優先範囲の外部であるならば、プロセスは調整されない。プロセスに対する現在のメモリ割当てが前のメモリの読みに等しくないならば、プロセスは調整されない。またもしもプロセス状態がそれが排気されたことを示すならば、調整は必要ない。他方、これらの基準に合致しないプロセスは調整が適当であるものとして設定される。

プロセスに対する調整を行うために、調整装置22によつて使用されるコンピュータシステム制御ルーチンはDLB30によつて呼出される。これは一般にプログラムを走行させることを要求するプロセスに非同期的にトラップを送ることによつて行われる。走行のためにプロセスに与えられるプ

装置24に転送され、sysadjustルーチン中で使用され、プロセス動作セットに対する制御値を置く。このようにしてプロセスは調整され、それによつてDLB30によつて制御される値を動作システムが供給される。これはDLB30がシステム情報に基づいて各個々のプロセスに対して調整を行わせることを可能にする。

この点において、LAST-MEME、MEM-AMT、PRE-ADJUSTS、およびPF-SINCE-ADJパラメータは更新され、プロセス特性パラメータ情報は蓄積される。計算可能なプロセスの数は基準に対して蓄積され、LAST-BALSETの値は更新されてDLBプロセス、ゼロプロセス、およびリストからの頁交換プロセスに移行し、調整されるべきプロセスを正確に反射する。

一度プロセスWSが調整されると、DLB30は一般的要求またはコンピュータシステムにおける負荷にしたかがつて新しいシステム動作パラメータを設定する。計算可能なプロセスの数は1以下

であるかどうかを検出するために比較される。これが真であれば、量子の値は表IIに示されたように量子ベースパラメータ下に蓄積された値に設定される。他方量子に対して選択された値は2または現在の計算可能なプロセスの数によつて計算された量子ベース値の商の2倍の大きいほうによつて決定される。

CPUの形式が特定された形式に一致する場合には、量子に対して割当てられた値は対応して調整される。これは、CPU時間およびメモリ割当てがより高い処理速度の利点を得るように調整されることができるよう何らかの形式で遭遇する処理速度差を考慮して使用される。所定のコンピュータシステムCPUが十分に高いレベルで行う係数はよく知られている。

A W S T I M Eに対する値は量子または5のいずれかに対して最大に等しく設定される。P I X S C A Nは最後の平衡プラス5または依然として最低である量子、または現在の平衡セットまたは10のいずれか最大のものに設定される。

とき値を重ね書きすることによつてD L B 30を設けることによつて行われる。

パラメータに対する変更の付加を開始する前、またはプロセスNを調整する前のD L B 30が休止される期間は2つのファクターによつて決定される。第1はコンピュータシステムが動作システムパラメータの最後の変更に通切に再調整するために必要な平均時間の長さおよび所望および実際のシステム特性間の偏差の程度である。

第1のファクターは平均の大型多使用者コンピュータシステムに対しては約6秒程度であることが認められた。明らかに当業者は考慮されるコンピュータシステムによつて経験される平均付加にしたがつてこの数が増加することが容易に理解されるであろう。しかしながら、経験では6秒はこの発明の方法に対して良好な最小サンプリング速度または時間を与えることが示される。

6秒毎に、D L B 30はプロセス変化を開始すべきか、待機すべきかを発見するためにコンピュータシステムの迅速な検査またはチェックを行う。

I O T Aパラメータは1に設定される。行われ、多数の調整はインクレメントされ、システムは将来の調整を行うために歴史的または統計的ベースを形成するためにシステム情報を集めるために前進する。

それからD L Bはコンピュータシステムが変化を行い、それから上述の待機モードへ進ことを確実にするために予め定められた短い遅延に対して待機する。

この発明の方法における次のステップは、設定されている新しいパラメータ値でコンピュータシステムが動作することができるよう予め定められた期間にわたつて休止することである。コンピュータシステムは一般的に欠陥初期化値を使用する処理を開始するけれども、新しい値のセットは任意の時に実行されることができるとは当業者には容易に理解されるであろう。すなわち、動作パラメータに対する新しい初期化値はパワーアップまたはリセットにおいてシステムスタートの一部として自動的に呼出し、或いは初期化が生じる

6秒の点で、コンピュータシステムにおける全体のメモリ量は蓄積されて現在の自由メモリと比較される。もしも自由メモリが全体のメモリの10%以下であれば、D B D L B 30はそれ自身を“パニックモード”に設定し、直ちにrでプロセス調整に戻る動作を開始する。もしも自由メモリがメモリ全体の50%以下であれば、D L B 30はそれ自身を“ウオーリーモード”に設定し、ウオーリー待機パラメータによつて指定された予め定められた期間待機を続ける。もしも自由メモリがメモリ全体の50%より大きければ、D L B 30は“正常”モードに設定され、正常待機パラメータによつて指定された期間待機を続ける。各程のモードはプロセス調整の前にD L Bが待機する時間を決定する。

モードテストが使用されて動作システム要求における即時注意対象々の変化を要求する負荷における突然の変化に適応させる。これはコンピュータシステム高い要求または迅速に変化するダイナミック部かによつてさえも応答できることを確実にする。

好ましい実施例において、3つの異なるサンプリング速度または待機周期として実行される応答の3つのレベルを設定することが有効であることが認められた。第1の速度は6秒の速度であり、それは初期化において、システムが“パニック”モードに設定されるときに使用される。第2のレベルは約30秒に設定され、“ウォーリー”モード待機として処理される。このモードでは所望の特性において顕著な変化があるが、システムは徐々に負荷に適合する。第3のレベルは90秒程度であり、“正常”モード待機を表わす。ここでは追加的または変化のない特性が所望され、システムは比較的長い動作周期にわたって少量の特性の変動を生じる。表Ⅱに示すようにこれらの予め定められた待機期間は負荷平衡装置30により使用される方法の初期化において負荷される。

当業者にはコンピュータシステムのリソースおよび負荷におけるダイナミック変化がそれを有用と認める他の待機期間が使用可能であることが容易に理解されるであろう。

ユータに対する応答性の増加は明瞭に達成され、システムロックアウトは低い優先度のプロセスにおいても実質的に除去される。

この発明の方法はプロセスの周期的調整を使用し、間に待機期間を有する。すなわち、この方法は全てのプロセスを連続的にチェックするのではなく、大きな要求をCPU時間に置くように調整を行うものである。

それ故、内部メモリおよび中央プロセッサ単位時間のようなコンピュータリソースに与えられた負荷プロセスを最良にするように多使用者コンピュータ環境におけるプロセスのためのシステム動作パラメータをダイナミックに自動的に調整する新しい方法および装置が説明された。さらに説明した方法はリソース割当てを最良にするための特性値の調整を含んでいる。

他のパラメータ表および値がこの発明の方法で使用できることを理解すべきである。

以上好ましい実施例の説明は説明の目的で記載されたものである。これはこの発明の全てを記載

この発明の方法は、18メガバイトの実質メモリ容量を有するVAX/VMS動作システムを使用するデジタルエキップメント社のVAX11/785型コンピュータシステムによつてじっこひうされる。システム同調を使用するにもかかわらず、システムはしばしば2000~3000頁の自由メモリを有するに過ぎない。この発明の方法を行った後、自由頁リストは平均11000乃至12000頁に増加した。

この発明の方法はまた他のデジタルエキップメント社のVAX/VMS型コンピュータシステムで実行することができ、それはI/Oボトルネックにより長い処理遅延および多くのシステムロックアウトを経験する。このコンピュータシステムに対するプロセス要求に回答する平均応答時間は第3図にライン80として示され、4時間の期間にわたっている。この発明の負荷平衡装置はコンピュータシステムおよびシステム応答に対して累積された統計的データにおけるアクティブプロセスとして設定された。この発明を使用する平均システム特性第3図にライン90で示されている。コンピ

したのではなく、またこの発明を説明された正確な形態のみに限定するものでもない。上記の説明から多くの要形変更が可能であろう。ここで説明した実施例はこの発明の原理を最もよく説明するためのものであり、その原理を適用することにより多くの実施態様が可能である。したがってこの発明の技術的範囲は特許請求の範囲の記載によつてのみ限定されるべきものである。

表 III

パラメータ	初期値	単位
HORN-WAIT	30	秒
HORRY-WAIT	30	秒
PAHIC-WAIT	4	秒
PAHIC-MEM	2000	頁
MOHUY-MEM	5000	頁
HO-ADJ-BASE	5000	頁
HO-ADJ-INCR	200	頁
QUANTUM-BASE	14	(1/256秒)
ABANDONED	120	
MIN-PRIOR	1	(レベル)
MAX-PRIOR	16	(レベル)
MEM-MIN	110	頁
ABS-MIN	100	頁
ADJ-MIN	20	頁
CPU-MIN	3	
MAX-FAULTS	10	(10秒)
MIN-MEM-ADJ	20	頁
MAX-MEM-ADJ	150	頁
DAS-MEM-ADJ	60	頁
DLB-MSDECR	180	頁
DLB-MSINCR	503	頁
SYS-GROUP	1	

表 IV

外部システムパラメータ	初期値	単位
HPW-HILIMIT	2500	又は全メモリ頁の15%
HPW-LOLIMIT	1250	又は全メモリ頁の50%
HPW-HAITLEMIT	16000	又は全メモリ頁の50%
SYSHWCHT	1000	
WSMAX	4000	
IRPCOUNT	300	
IRPCOUNTV	1000	
LRPCOUNT	50	
LRPCOUNTV	80	
SRPCOUNT	900	
SRPCOUNTV	1000	
HPAGEDYN	500000	
PAGEDYN	500000	
HPAGEVIR	500000	
VIRTUALPAGECNT	15000	
PAGEFIEL.SYS	1	物理的メモリ頁の2倍
SWAPFILE	1	交換ファイルのWSMAXより大きい場合は
HSEXTENT	4000	or WSMAX

表 V

ダイナミックプロセス調整: 調整
調整のためのプロセス調整チェック、プロセス特性チェック、および
プロセスリソースの調整
実行
リサイクルカウンタ1インクリメント
アクチバカカウンタおよびCOMカカウンタ調整
全メモリおよび自由メモリ解放
自由メモリを調整なし ならば
実行 調整なし
終了
その例
例1: アクチバカカウンタ=最大の1または
(LAST RAISFT-LAST ABANDONED)
例2: MEM SAFE AMOUNT=MEM MIN+
FREE MEM/IMAXの1または0.70 (アクチバカカウンタ)
例3: 例子=2
例4: IOTA=2
終了
プロセス実行
カウンタ初期化、予備調整、LAST RAISFT、およびLAST ABANDONEDの
例1のプロセスによりスタート
実行
プロセス状態チェック
それ以上プロセスがないならば
終了
プロセスが交換され、または休止されるならば
調整なし
その例

プロセス調整
プロセスインデックス調整
プロセス調整を始めるために現在の値をリセット
現在のプロセスID=例のプロセスIDでなければ
IDは新しい値を行なう。

終了
例の値として保持されたプロセス値を保持
調整データリセット実行
調整のMEM AMOUNTをRAS MEM AMOUNTに設定
調整のMEMを0に等しく設定
ADJ=0であるからPF設定
レベルを0に設定
終了
終了
調整が完了すれば、LAST RAISFTを1インクリメント
プロセスが調整でなければLAST ABANDONEDを1インクリメント
プロセス状態がCOMに等ければ、
カウンタを1インクリメント
ADJおよびADJ=0またはSYSグループにない 実行ならば
プロセスチェック実行
プロセスIDおよびインデックス設定
現在のIDにより調整されるからPFインクリメント
現在のメモリをABS PROC MINまたは
調整が0でなく、前のCPUをCPU MINから
現在のCPUに変化するならば、例は、
現在の優先度が範囲の外であるならば、例は
現在のメモリが例のメモリに等しくないならば、
例の現在の値に等しく設定
レベルをLEVELに設定
終了

その他
レベルを減少ならば退出

終了

その他

このプロセスのために動作セット開始

設定: CPU AMNT=現在CPU-側のCPU

PF SINCE ADJ.1>最大値であるから

MEMAMT=LAST MFM AMT/2に等しい

終了

その間 MFMAMT=1.5 * LAST MFM AMT

終了

MFMAMT= (MAX MFM ADJ.または0) のMIN
(MIN MFM ADJ. MEMEAMT)

終了

INFOリセット実行

レベルが閾値より大きければ、

ADJ=0に調整

その他

保存のメモリを安全メモリならば復旧

MEMAMT=LAST MFM AMTおよび

現在のメモリ=LAST MEM

レベルはMAXに等しい (レベル=1, 0)

PF SINCE ADJ=0

復旧

終了

ADJAMT= (MEMAMT+現在WS-現在MFM)

終了

OUTSWAPPED プロセスレベル=ARANDONFに設定

初期状態プロセスレベル=MAX (レベル=1, 0)

実行

設定: PIXSCAN= (LAST RAISEのMINまたは
5 * 原子ベース/原子またはRAISEまたは10の最大
設定: IOTA=1

終了

TOT ADJUSTSをPRE ADJUSTSによりインクリメント

終了

STATS INFO取得

DIRプロセスサイクルSIN進行の取得

MFM SAFF AMTから各アクティブプロセスに格納されたメモリを
取得

TOT ADJUSTSから調整されたプロセスの全体数取得

PRE ADJUSTSから調整されたプロセスLASTの数取得

LAST RAISE=LAST ARANDONF取得

自由リストOFF エントリーの最大数取得

情報実行

全体メモリおよび自由メモリの取得

初期モードを故障モードに等しく設定

自由メモリが全体のメモリまたはパニックメモリの10%以下ならば、

パニックモード選択

自由メモリが全体のメモリの50%以下またはウォーリメモリ以下ならば

ウォーリモード選択

その他

正常モード選択

終了

モード=モード値 (6, 30, 90) に設定

実行

モードを1からモード値/パニック値1にインクリメント

パニック値により監視

自由および全メモリ取得

モードが初期モードよりも少ないならば退出

ADJAMTおよびDLB WSINCR値を調整装置へ送る
SYSSADJWSL実行

終了

状態が異なりならば退出

PRE ADJUSTSを1インクリメント

LAST MFM AMT=MEMAMT

設定: LAST MFM=現在のMFM

設定: PF SINCE ADJ=0

JOB INFOストローブ

終了

設定: 計算可能なプロセス-C カウント

設定: LAST BALANCE SET=LAST BALANCEの最小
またはRAISECNT

設定: LAST BALANCE SET=1またはLAST BALSET
の最大-3

パラメータ変更実行

システムに使用されたCPUの形式取得

システムにおける計算可能なプロセスの数取得

LAST RAISECNT取得

計算可能なプロセスの数≤1ならば

原子=原子ベースに設定

終了

その他

設定: 原子=2または2 (原子ベース/COMカウント) の最大

終了

CPUの形式が予定の型と一致していれば

調整値により多項式調整

終了

設定: AWSTIME=原子または5の最大

4. 図面の簡単な説明

第1図はこの発明の1実施例に使用するコンピュータシステムの制御ブロック図であり、第2図は、この発明の1実施例のフローチャートである。第3図は、コンピュータシステムの1例における平均システム応答時間を示す。

12…動作システムパラメータ制御装置、14…プロセス制御ブロック、16…WS制御装置、18…累積レジスタ、20…モニタ、24…プロセス調整装置、30…ダイナミック負荷平衡装置。

出願人代理人 弁理士 鈴江武彦

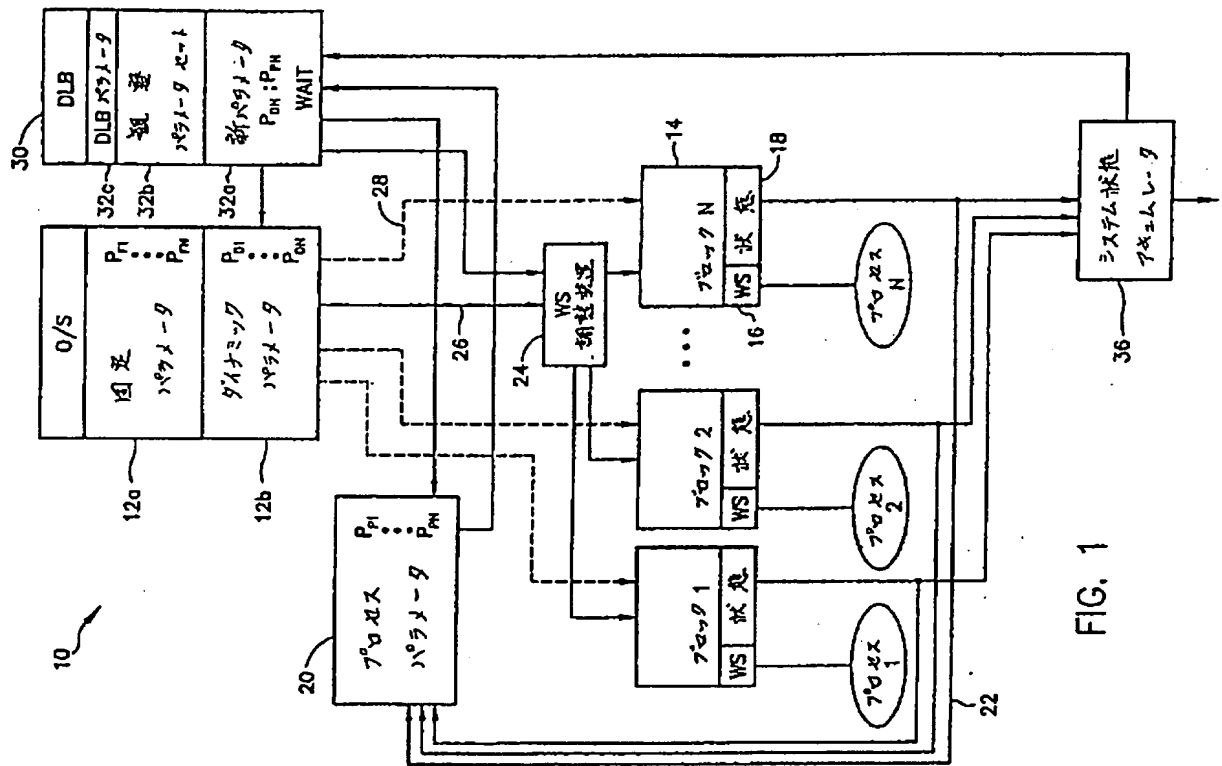


FIG. 1

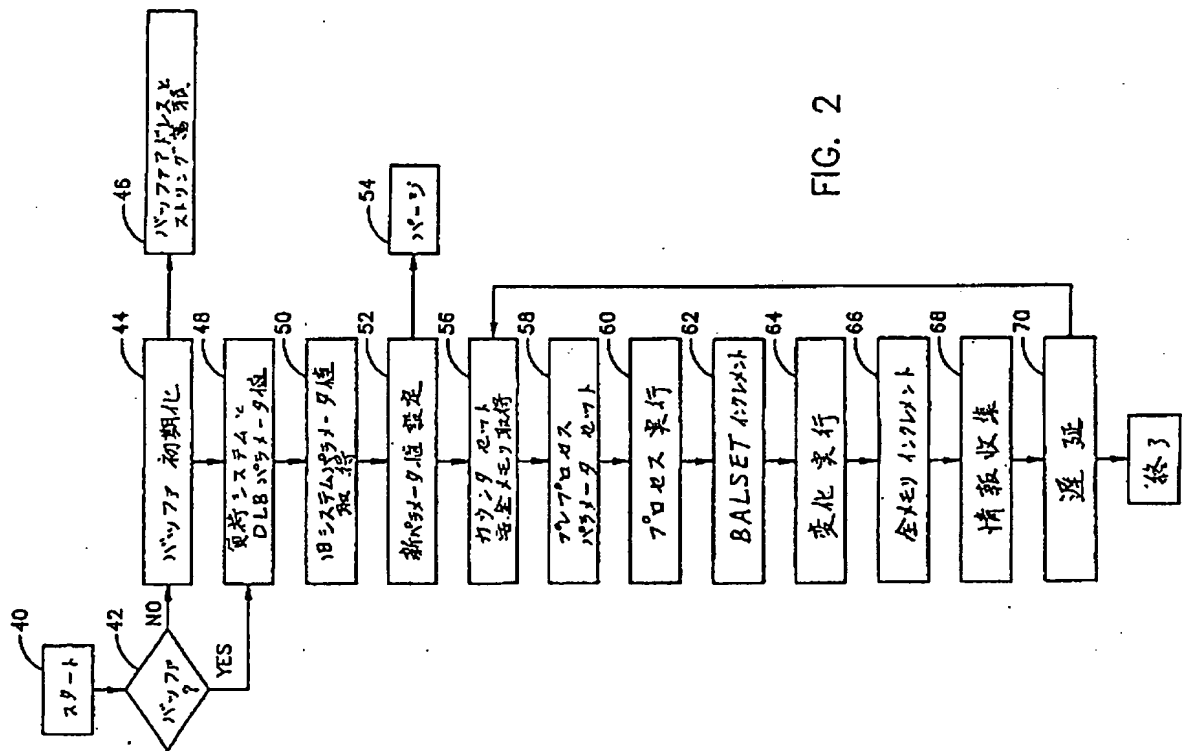


FIG. 2

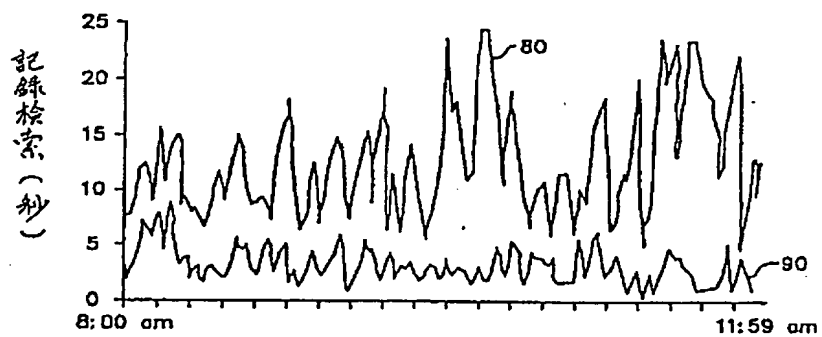


FIG. 3